Science of Information, Computation and Fusion

05 MAR 2012

Tristan Nguyen
Program Manager
AFOSR/RSL
Air Force Research Laboratory

Integrity ★ Service ★ Excellence

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **05 MAR 2012**	2. REPORT TYPE	3. DATES COVERED **00-00-2012 to 00-00-2012**
4. TITLE AND SUBTITLE **Science of Information,Computation And Fusion**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory,Wright-Patterson AFB,OH,45433-**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S)
		11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the Air Force Office of Scientific Research (AFOSR) Spring Review Arlington, VA 5 through 9 March, 2012

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**28**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

2012 AFOSR SPRING REVIEW

NAME: Science of Information, Computation and Fusion

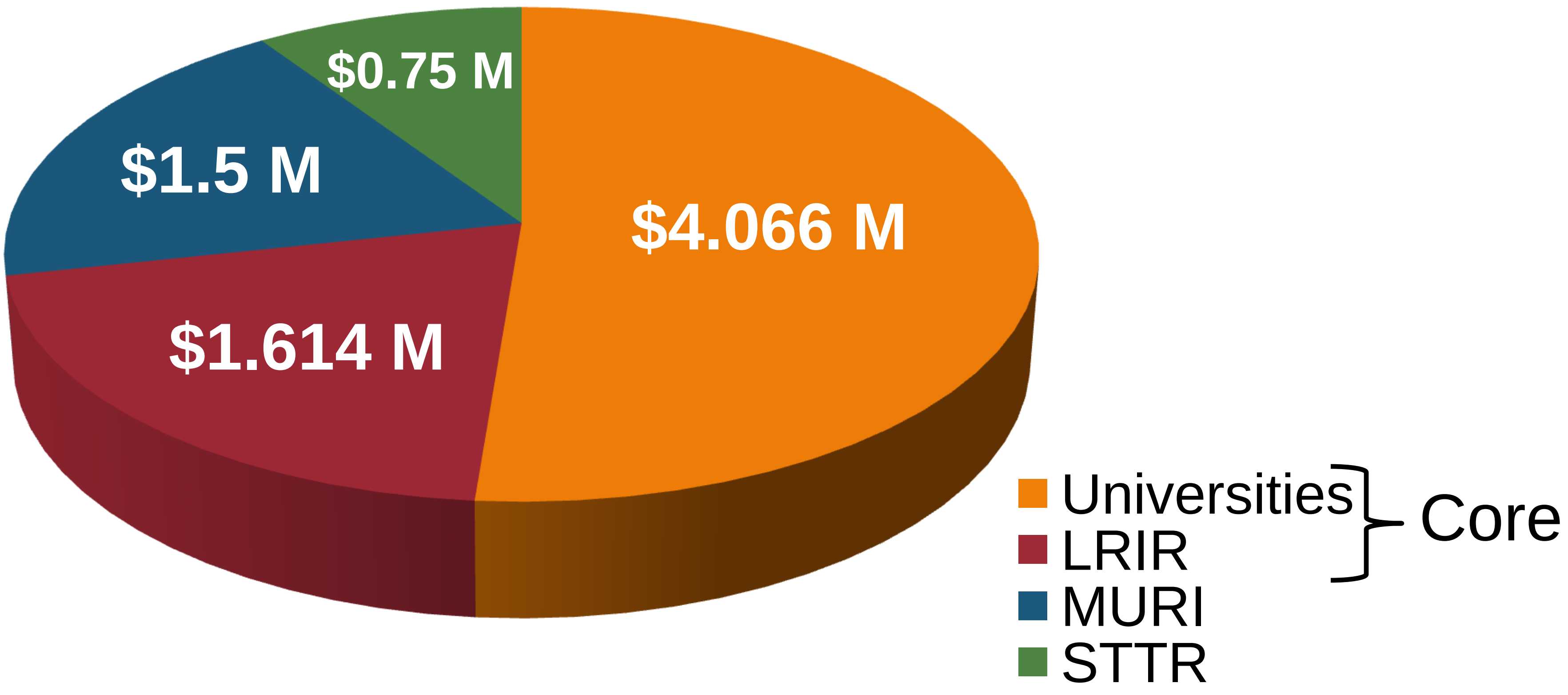
BRIEF DESCRIPTION OF PORTFOLIO:

Information fusion requires understanding of information across multiple domains. Challenges: (1) how to formalize domains of knowledge and information – structures and their interconnection; (2) how to mechanize patterns of reasoning in terms of computation.

LIST SUB-AREAS IN PORTFOLIO:

•Sub-Areas	Objectives
Data Discovery Information Science	• Discover structures in data and shape them into information • Methods for fusing heterogeneous data • Higher-level information and computation
Algorithmic Development Statistical Modeling	• Exploit structures • Leverage *a priori* models } To improve processing
Sensors & Radars Communication Networks	• Fusion of sensors and fusion of network data for detection • Performance of sensor fusion

Portfolio (FY12)



DISTRIBUTION A: Approved for public release; distribution is unlimited.

Air Force Relevance

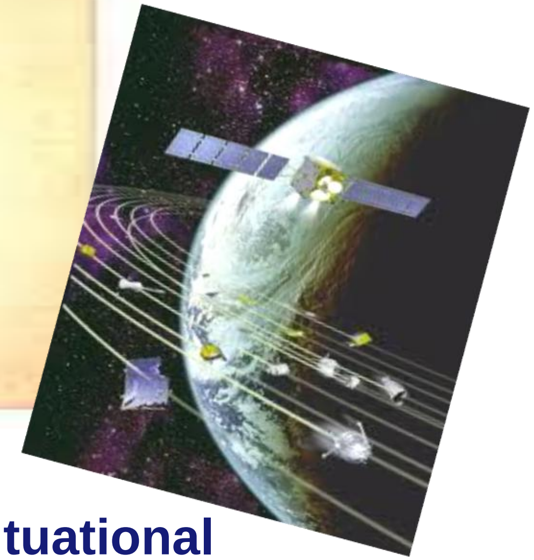
Autonomy

Cyber Domain

Program Contribution:

- Information Triage
- Automated Reasoning
- Human-Machine Interface
- Formal Verification

Intelligence, Surveillance, Reconnaissance



Space Situational Awareness

AFRL

Collaborations & Transitions

- **OSD/ONR** – Data to Decision
- **OSD/AFRL/RH** – Autonomy
- **DARPA/AFRL/RY** – Mathematics of Sensing, Exploitation and Execution (MSEE)
- **ARO's MURIs**:
  - *Multivariate Heavy-tailed Statistic*
  - *Revolutionizing High-dimensional Microbial Data Integration*
- Topological techniques for data analytics transitioned to Army's Phase-II SBIR.
- Compressed-sensing techniques transitioned to Trusted Access Program Office (TAPO).

Future Direction

Sub-Areas	Objectives
Data Discovery **Information Science**	• Discover structures in data and shape them into information • Methods for fusing heterogeneous data • **Higher-level information and computation**
Algorithmic Development **Statistical Modeling**	• Exploit structures • Leverage a priori models
Sensors & Radars **Communication Networks**	• **Fusion of sensors and fusion of network data for detection** • **Performance of sensor fusion**

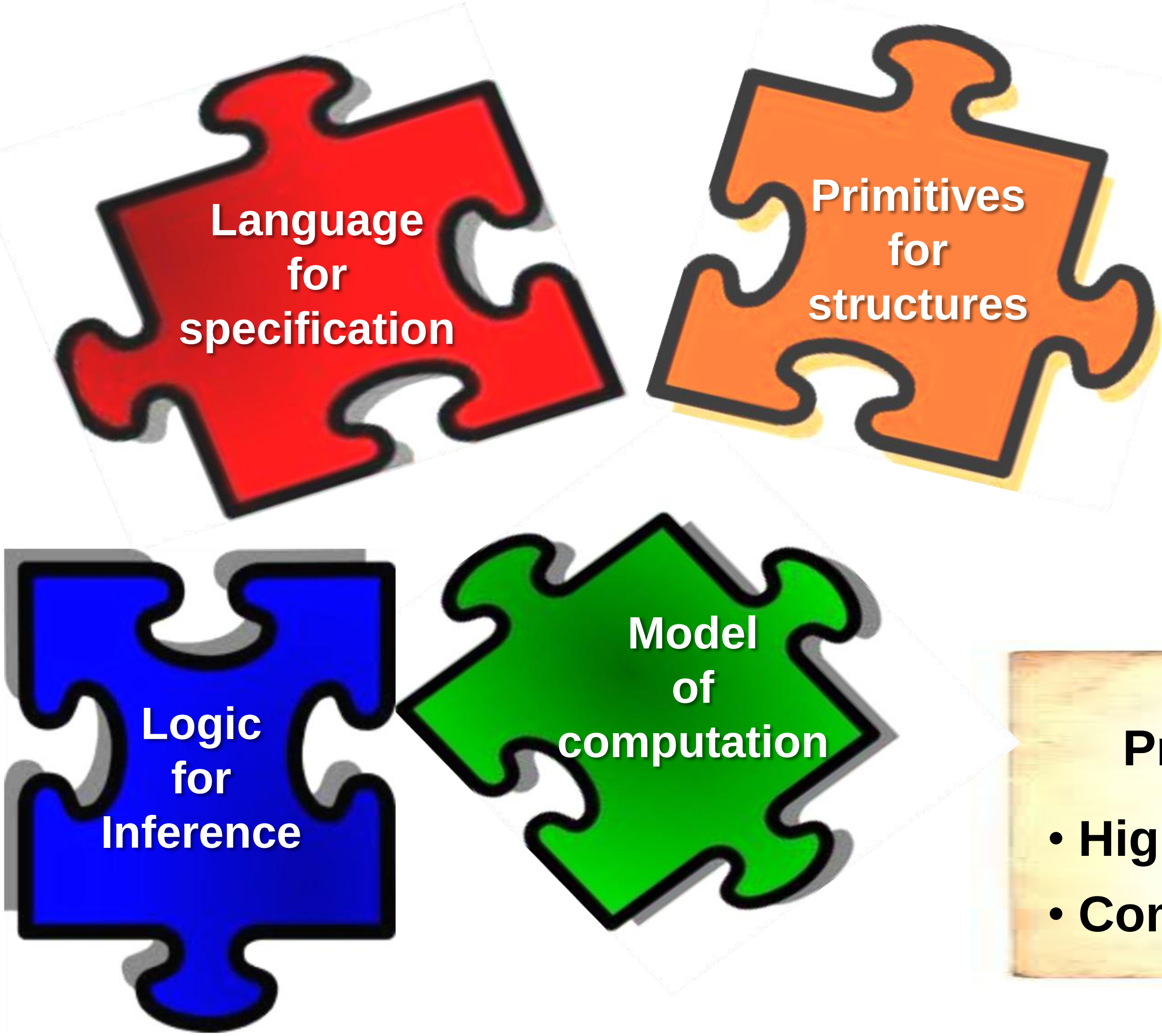
Emphasized

De-emphasized

Rationales:

- Avoidance of duplication
- Lack of attention to higher-level information and computation grounded on basic research
- Emerging scientific developments in this research area

Higher-level Information & Computation: Ingredients and Recipes?



Generic Programming

Constructive Mathematics

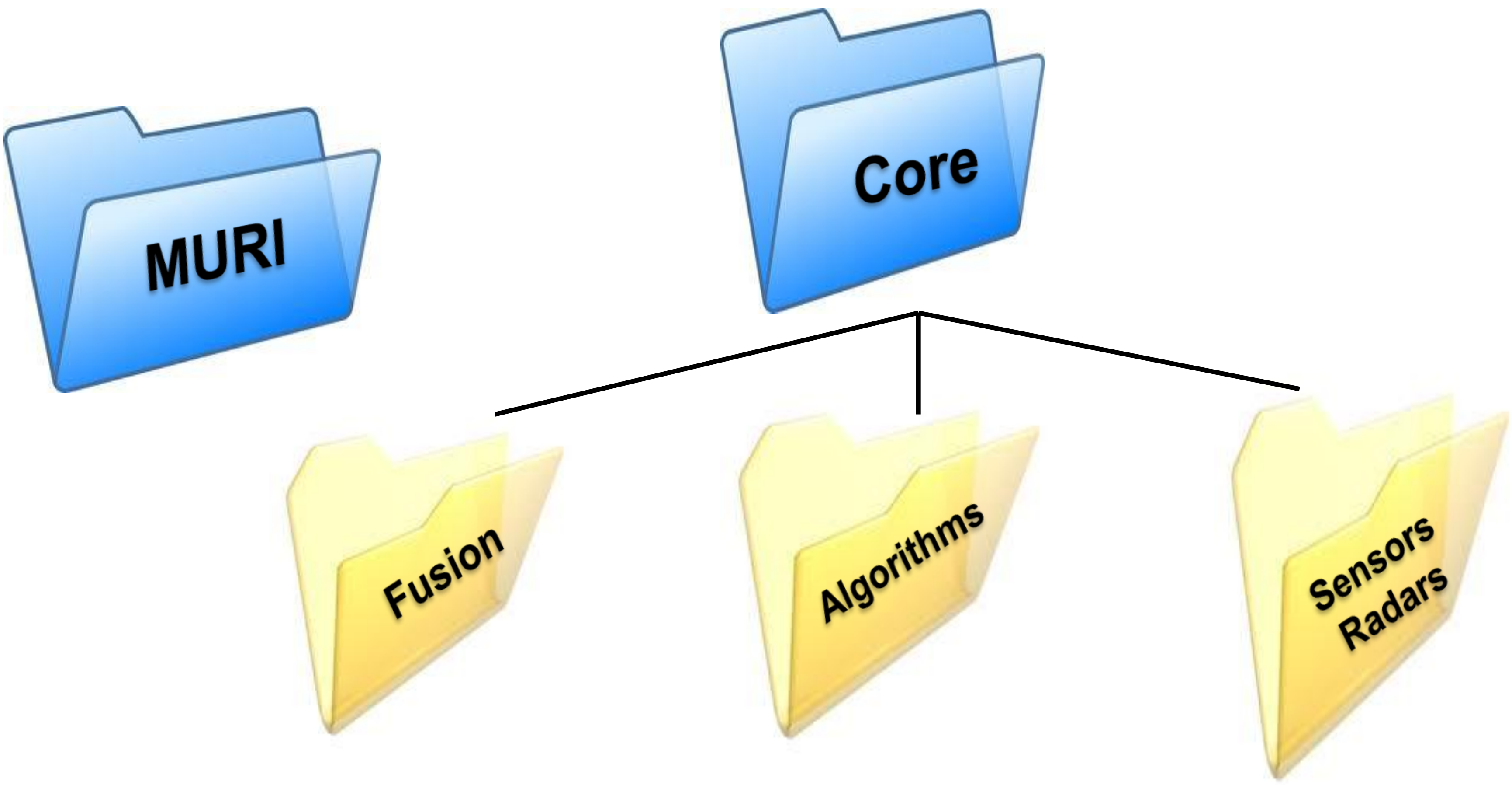
Cognitive Psychology

Program Challenges:

- **Higher-order Structures**
- **Computability**

AFRL

Highlights of Research Projects



Highlights of Research Projects

FY10 MURI - Control of Information Collection and Fusion

Penn
University of Pennsylvania
- Jadbabaie
- Koditschek
- Kumar
- Ribeiro

Berkeley
- Ramachandran
- Sastry
- Tomlin

Illinois
- Baryshnikov

Minnesota
- Giannakis
- Roumeliotis

Melbourne
- Howard
- Moran

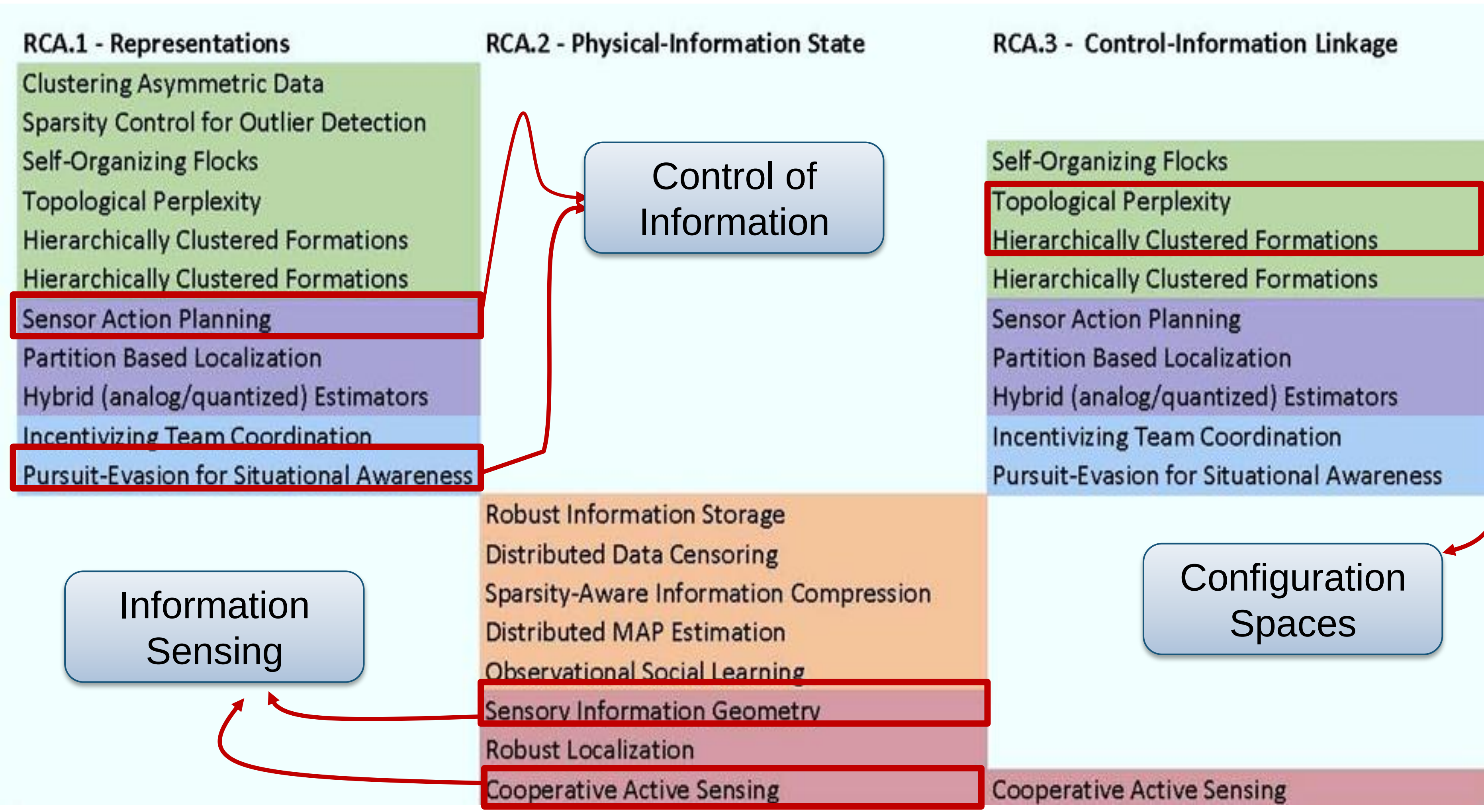
Objective: To formulate a new perspective on the *joint control* of *heterogeneous* information sources to simultaneously achieve quantified *informational* and *physical* objective.

Related MURI Topics of ARO:
- *Opportunistic Sensing*
- *Value of Information*

AFRL/RY's Involvement:
- Connection with the LRIR project on Layered Sensing
- Interaction between MURI team and RY's technical staff
- Agent for DARPA/DSO's MSEE

In Connection with AFOSR Programs:
- Intersection with all three sub-areas
- Dynamics and Control
- Sensing, Surveillance, and Navigation

FY10 MURI - Control of Information Collection and Fusion

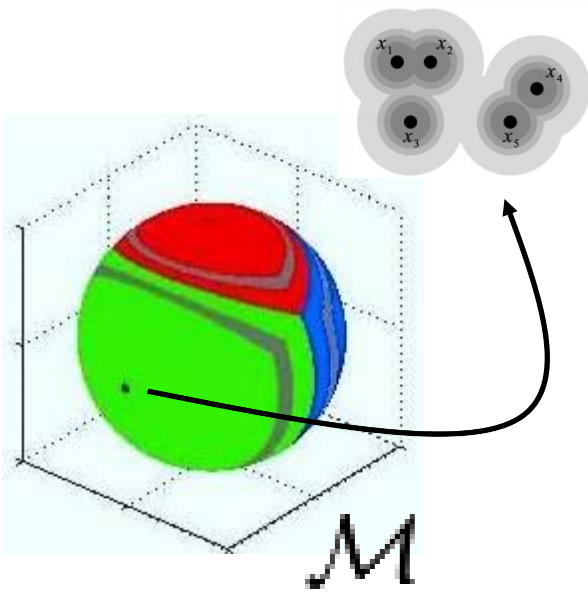


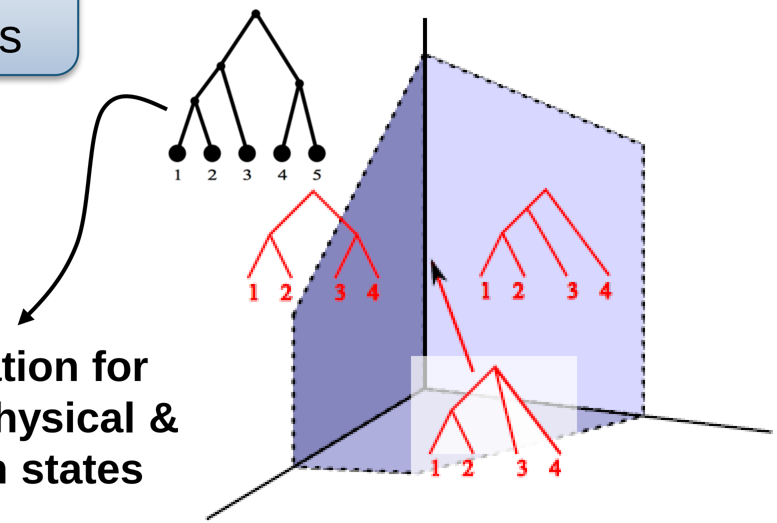
FY10 MURI - Control of Information Collection and Fusion

Space of Configurations

Space of Metrics Between Sensors

Space of Rooted Trees



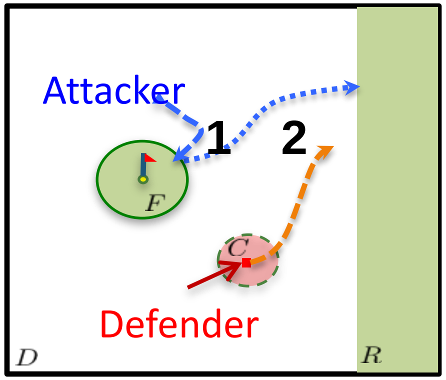


Representation for combining physical & information states

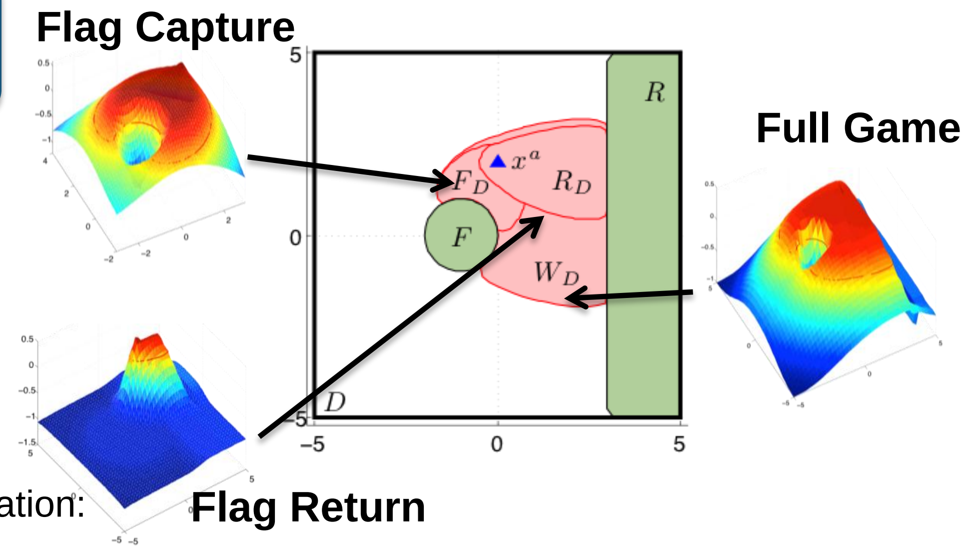
Goals:

- Investigate the (higher-order) structures of these spaces.
- Formulate control theory on these spaces.
- Formulate information sensing on these spaces.

FY10 MURI - Control of Information Collection and Fusion



Control of Information



Pursuit-Evasion Game

Initial Condition: $J(x,0) = \phi_R$

Modified Hamilton-Jacobi-Isaacs Equation:

$$\frac{\partial J(x,t)}{\partial t} + \min[0, H^*(x, \frac{\partial J(x,t)}{\partial x})] = 0$$

Subject to: $J(x,t) \geq -\phi_A$

Next: Increasing Game Complexity

- Multiple defenders versus attacker.
- More complex, interactive strategies between the attacker & defender.
- More complicated attacker-defender configurations.

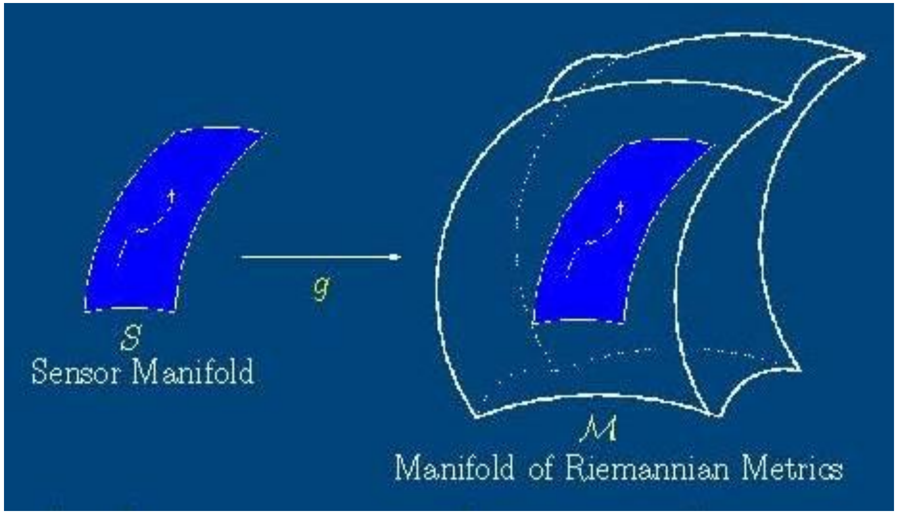
FY10 MURI - Control of Information Collection and Fusion

Information Sensing

Model of Sensor $\left\{ f(X;\theta) \right\}$ **Parameterized by a configuration space** S

Sensor Information $(\mathcal{I}(\theta))_{i,j} = \mathrm{E}\left[\frac{\partial}{\partial\theta_i}\ln f(X;\theta)\frac{\partial}{\partial\theta_j}\ln f(X;\theta)\right]$

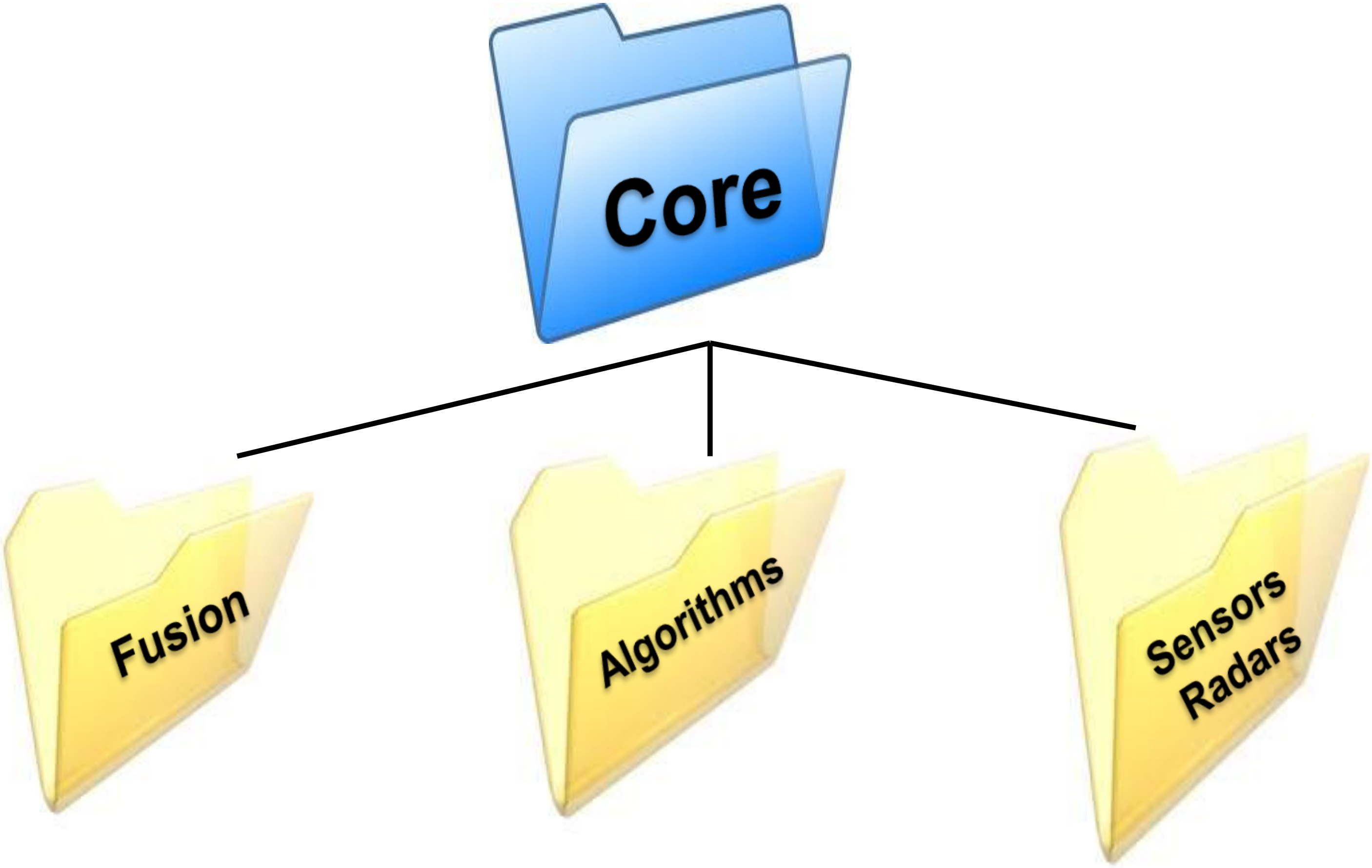
Idea: each sensor configuration gives a Riemannian metric on S



$\mathcal{M}$ **: infinite Riemannian manifold if** S **is a manifold**

Goal: Study path planning on ***S*** using structures on $\mathcal{M}$

Highlights of Research Projects



Academic Projects in Core Program

•Sub-Areas	Objectives
Data Discovery **Information Science**	• **Discover structures in data and shape them into information** • **Methods for fusing heterogeneous data** • **Higher-level information and computation**
Algorithmic Development **Statistical Modeling**	• **Exploit structures** • **Leverage *a priori* models** **To improve processing**
Sensors & Radars **Communication Networks**	• **Fusion of sensors and fusion of network data for detection** • **Performance of sensor fusion**

Some PIs:

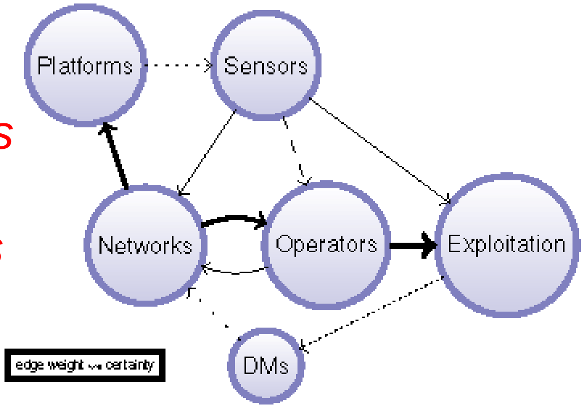
- J. Culbertson (AFRL/RY)
- G. Carlsson (Stanford U.)
- C. Priebe (Johns Hopkins U.) & D. Marchette (NSWC-Dahlgren)
- T. Klausutis, S. Butler, J. Curtis (AFRL/RW)
- D. Blei (Princeton U.)
- R. Devore (Texas A&M U.)
- A. Montanari (Stanford U.)
- A. Hero III, R. Raich (U. Michigan & Oregon State U.)

Category Theory & Layered Sensing

J. Culbertson et al., AFRL/RY

❑ Project begins in FY12.

❑ Layered Sensing aims to generate *situational awareness* from available resources (*sensors, platforms, networks, human operators, decision makers*) with required *attributes* (*persistent coverage, trusted sensing, information triage, adaptability, understanding social dimensions*)



❑ **Goals**:

- To model *heterogeneous* systems with uncertainty in a *universal* context and *layers* of abstraction.
- To study basic properties of these systems: composability, stochasticity, relations.
- To develop a *model for Layered Sensing* & use this model for *fusion*.

Technical Approach: Category of Probability of Spaces

Objects of Study

Probability spaces $(\Omega, \Sigma_\Omega, P)$ and maps between them

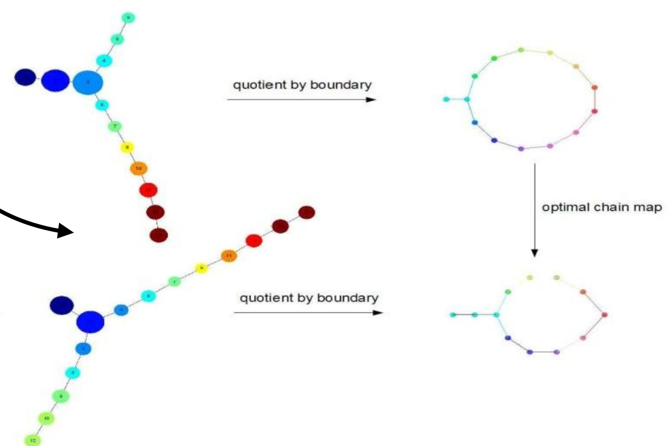
- Ω is the event space
- Σ_Ω is a σ-algebra on Ω
- $P : \Sigma_\Omega \to [0, 1]$ is a probability measure

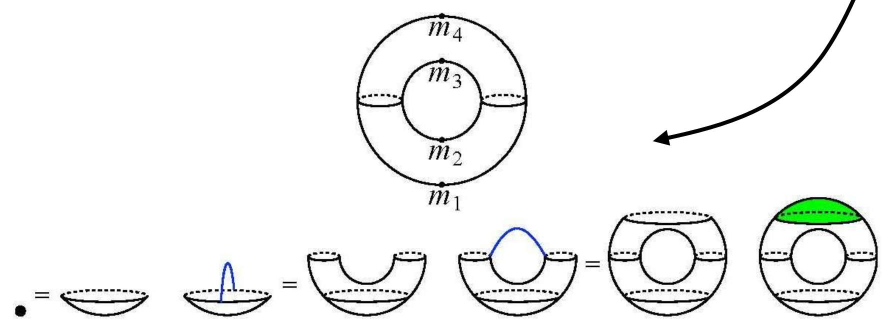
Data Fusion via Topology

G. Carlsson et al.

Current Tasks:

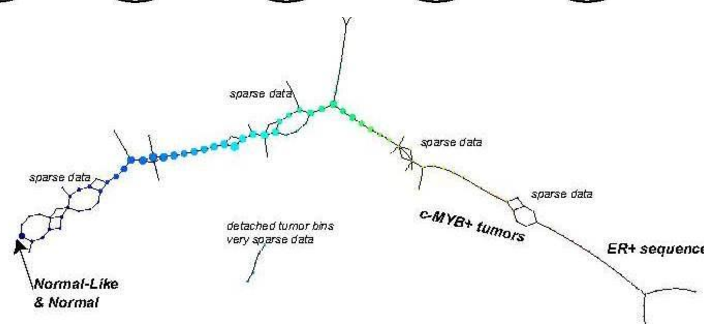
1. Data fusion: To study maps between simplicial complexes (models of topological spaces.
2. Information discovery: Morse-theoretic decomposition of spaces for finer analysis.
3. Sampling issue: To study how different samples affect the outcomes.





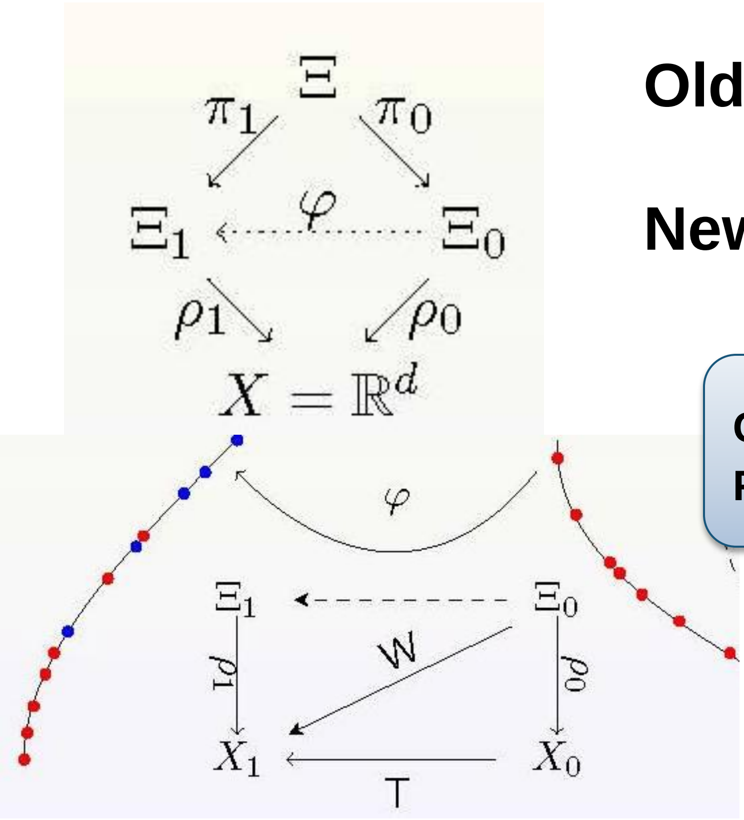
Application of Past Results: Identification Breast-Cancer Patients with Unique Mutational Profile & Excellent Survival

In *Science News Prime,* Oct. 3, 2011 & *PNAS* 2011



Data Fusion via Embedding & Matching

C. Priebe, D. Marchette

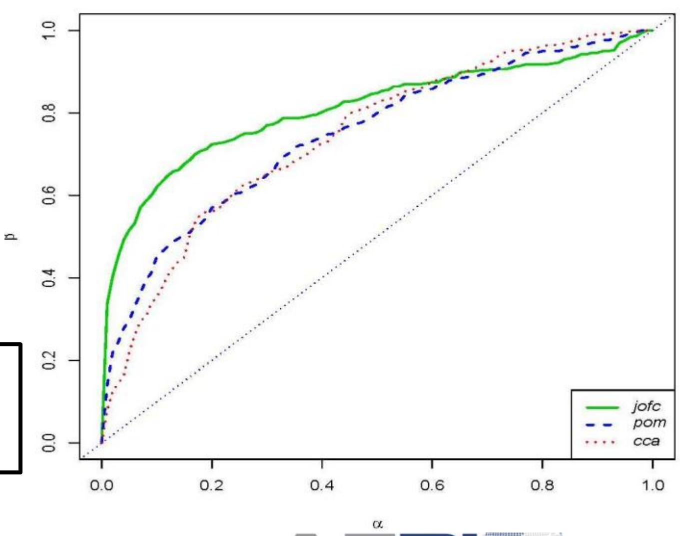


Old: Δ_0, Δ_1 are used to construct $\boldsymbol{\rho_0}$, $\boldsymbol{\rho_1}$

New: $$M = \begin{bmatrix} \Delta_0 & W \\ W^T & \Delta_1 \end{bmatrix}$$

Commensurability – *W* is used for matching
Fidelity – Δ is used for preserving dissimilarities

Classification of English/French Wikipedia documents



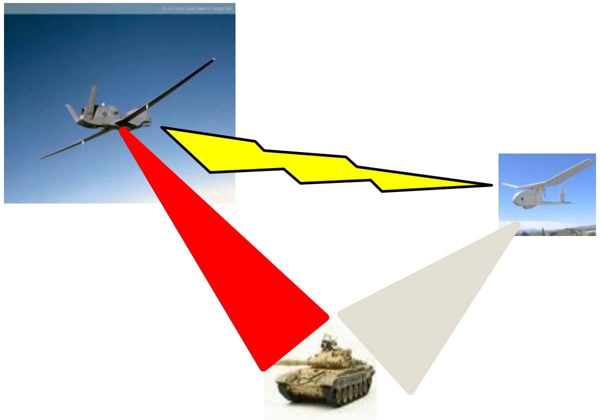
DISTRIBUTION A: Approved for public release; distribution is unlimited.

Manifold Embeddings & Target Acquisition

T. Klausutis, S. Butler, J. W. Curtis, AFRL/RW

Mission: Autonomous Target Acquisition

- **Find and Fix:** Target detection & identification
- **Track:** Tracking & aimpoint selection
- **Assess:** Battle damage indication and assessment



Technical Challenges:

- Multiple platforms coordinate to find and track targets
- Computational capabilities vary with platforms
- Platforms may have different sensor capabilities

Scientific Opportunities:

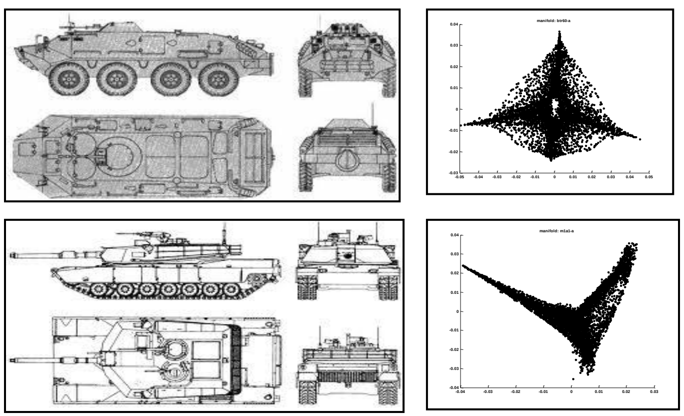
- Nonlinear manifold learning, interactive or supervised learning for classification.
- Low-dimensional data representation and analysis.
- Compressed sensing and information reconstruction.

Manifold Embeddings & Target Acquisition

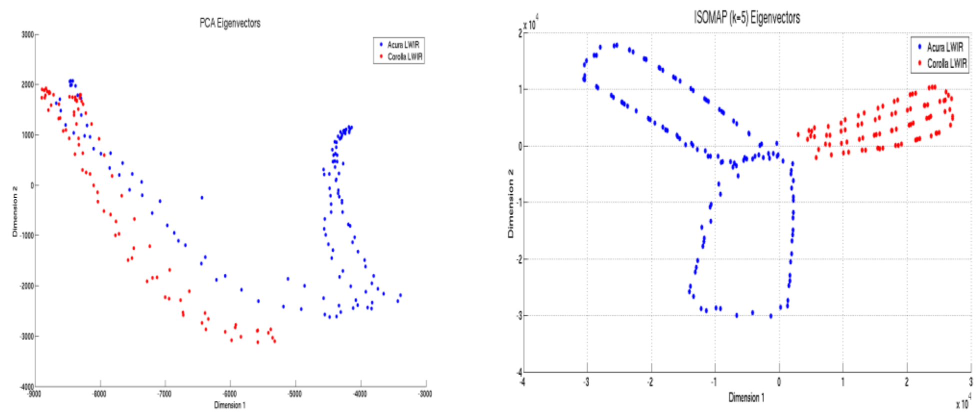
T. Klausutis, S. Butler, J. W. Curtis, AFRL/RW

Current Progress:

- Investigating nonlinear dimensionality reduction to reduce computational overhead.
- Comparing PCA and its nonlinear counterparts.
- Verifying the nonlinear techniques on different datasets.



81-D feature vectors reduced to 2-D
(Ladar Data)



Separation of Targets: PCA vs. ISOMAP
(Synthetic LWIR Data)

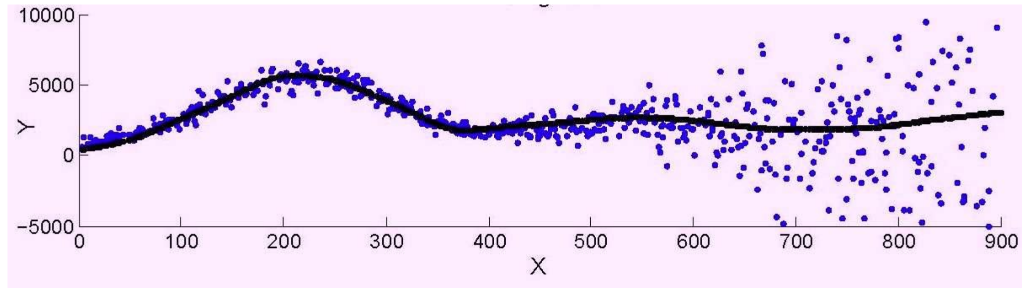
Next:

- Study sparse and low-dimensional machine learning in this context.
- Study algorithm’s robustness to variations in scale, intensity, etc.

Dirichlet Process Mixtures of Generalized Linear Models

D. Blei et al.

$$
\begin{aligned}
P &\sim DP(\alpha \mathbb{G}_0), \\
\theta = (\theta_{i,x}, \theta_{y,i}) | P &\sim P, \\
X_i | \theta_{i,x} &\sim f_x(\cdot | \theta_{i,x}), \\
Y_i | x_i, \theta_{i,y} &\sim GLM(\cdot | X_i, \theta_{i,y})
\end{aligned}
$$



Key Ideas:

- Local grouping of *X* and *Y* with varying mixture components.
- The number of groups need not be known in advance.
- *X* can simultaneously be *categorical* and *numerical* variables.
- A tool that can model many response types and many response shapes.

New Regression Technique

Recent research featured in The Economist

The Economist

World politics | Business & finance | Economics | Science & technology | Culture | The World in 2012 | Blogs | Debate | Multimedia | Print edition

Organising the web

The science of science

How to use the web to understand the way ideas evolve

Apr 28th 2011 | from the print edition

COMPUTER scientists have long tried to foist order on the explosion of data that is the internet. One obvious way is to group information by topic, but tagging it all comprehensively by hand is impossible. David Blei, of Princeton University, has therefore been trying to teach machines to do the job.

Princeton University

AFRL

High-dimensional Approximation

R. Devore et al.

Scientific Problem: How to recover a function in high dimension with small errors when only a finite number of evaluations are known?

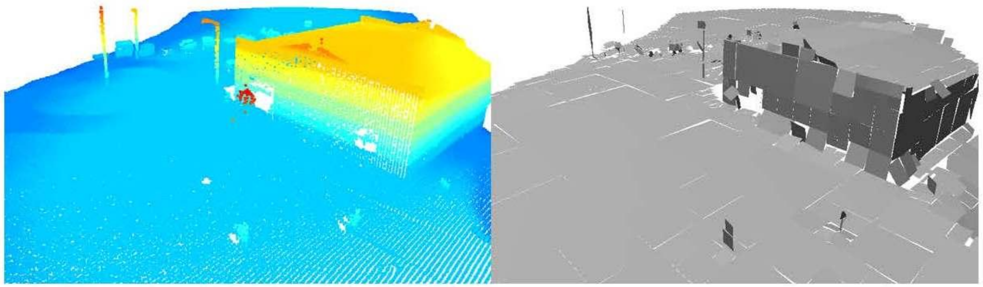
Assumptions: (1) some form of sparsity
(2) reduced number of independent variables.

Theorem (DeVore-Petrova-Wojtaszczyk) Suppose that

- $f \in C^s(\Omega)$
- There is a g and $j_1, \ldots, j_d$ (both unknown to us) such that $\|f(x_1, \ldots, x_D) - g(x_{j_1}, \ldots, x_{j_d})\|_{C(\Omega)} \le \epsilon$.
- By making $n = C(d, S)L^d \log_2 D$ non-adaptive point queries we can recover f by $\hat{f}$ to the accuracy

$$\|f - \hat{f}\|_{C(\Omega)} \le C(S, d)\{\|g\|_{C^s([0,1]^d)} L^{-s} + \epsilon\}$$

Another Project: Reconstruction, compression of terrain point cloud data using simple function approximation



Point Cloud

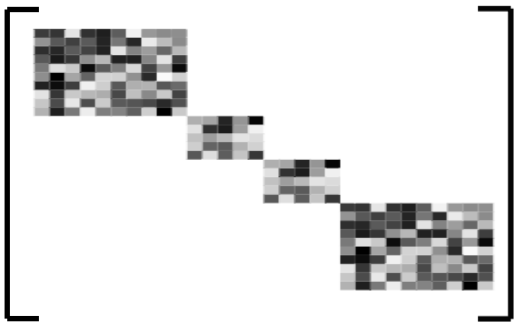
Reconstruction using **PLANE**, $\eta^* = 0.006$

New Framework for Compressed Sensing

A. Montanari et al.

New joint work with D. Donoho and A. Javanmard

$A =$



$, \quad y = Ax + w, \quad x \sim \rho_x$

***m*x*n* matrix**

Result - Suppose $x \sim \rho_x$. Sensing matrix A of this type can be constructed such that:

- If $y = Ax$ is observed, then x can be recovered with high probability under some hypotheses on m, n, and ρ_x.
- The recovery is robust to noise.

Notes: There are instances in which this new result supersedes the current one. Sparsity of **x** is replaced by $x \sim \rho_x$ and other hypotheses on x.

Sensor Fusion via Shannon/Rényi Entropy

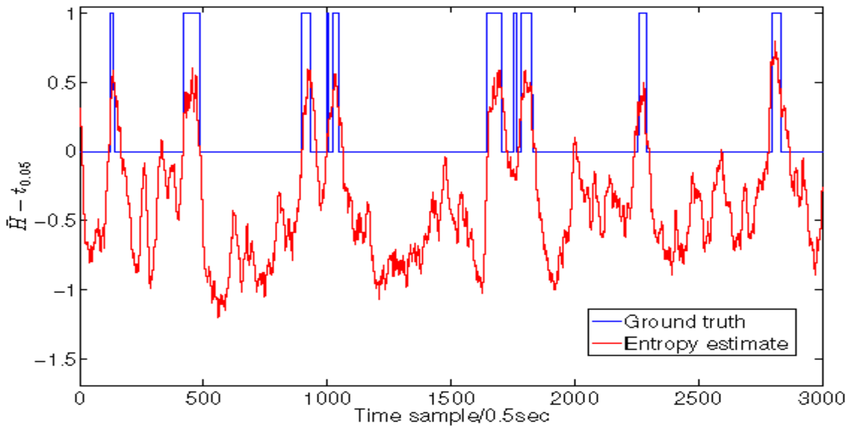
A. Hero III, R. Raich

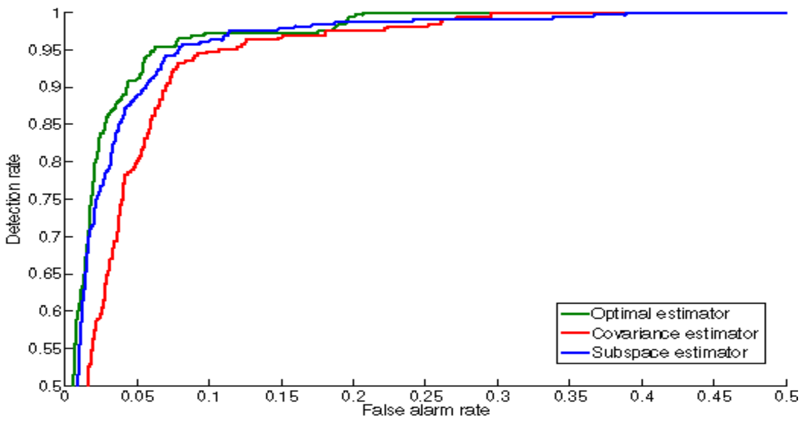
Motivation: Sensor fusion via Shannon's information theory requires computation of entropy or divergence. **Nuisance**: Joint probability density function is not known.

Entropy
$$G(f) = \int g(f(x))f(x)d\mu(x)$$

Estimator
$$\hat{\mathbf{G}}(f) = \frac{1}{N}\sum_{i=1}^{N} g(\hat{\mathbf{f}}(\mathbf{X}_i)).$$

i.i.d samples
$$\{\mathbf{X}_1, \ldots, \mathbf{X}_N, \mathbf{X}_{N+1}, \ldots, \mathbf{X}_{N+M}\}$$

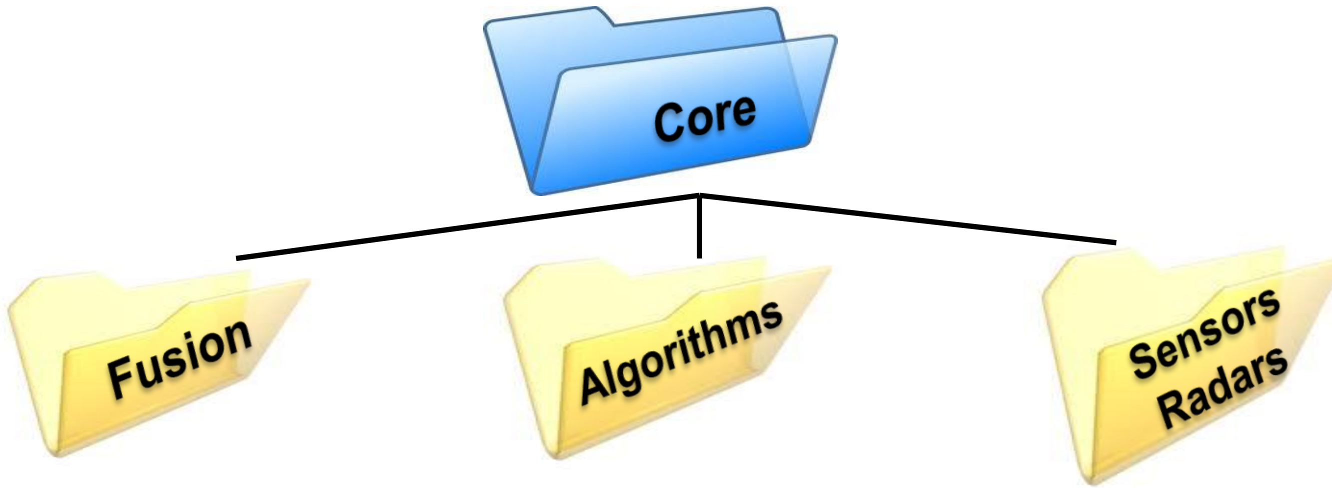




Result:
- Derive Variance & Bias
- Establish Gaussian central limit behavior
- Extend to more general cases

Summary

Current Portfolio



- To formalize and mechanize information-related concepts in a computational framework.
- Constant need to develop more techniques to deal with diverse types of information.
- Bridging the gap between humans and sensors in the information domain.

Thank You!